4 un 13 - 1

SEQUENCE LISTING

| (1) | GENERAL | INFORMATION: |
|-----|---------|--------------|
| | | |

5

(i) APPLICANT:

- (A) NAME: Akzo nobel n.v.
- (B) \$TREET: Velperweg 76
- (C) CITY: Arnhem

10

- (E) COUNTRY: The Netherlands
- (F) POSTAL CODE (ZIP): 6824 BM
- (G) TELEPHONE: 0412-666379
- (H) TELEFAX: 0412-650592
- (I) TELEX: 37503 akpha nl

15

- (ii) TITLE OF INVENTION: Novel estrogen receptor
- (iii) NUMBER OF SEQUENCES: 28

20

- (iv) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

25

30

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1434 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

| ATGAATTAC A | GCATTCCCAG | CAATGTCACT | AACTTGGAAG | GTGGGCCTGG | TCGGCAGACC | 60 |
|--------------------|------------|------------|------------|------------|------------|------|
| ACAAGCCCAA | ATGTGTTGTG | GCCAACACCT | GGGCACCTTT | CTCCTTTAGT | GGTCCATCGC | 120 |
| CAGTTATCAC | ATCTGTATGC | GGAACCTCAA | AAGAGTCCCT | GGTGTGAAGC | AAGATCGCTA | 180 |
| GAACACACCT | TACCTGTAAA | CAGAGAGACA | CTGAAAAGGA | AGGTTAGTGG | GAACCGTTGC | 240 |
| GCCAGCCCTG | TTACTGGTCC | AGGTTCAAAG | AGGGATGCTC | ACTTCTGCGC | TGTCTGCAGC | 300 |
| GATTACGCAT | CGGGATATCA | CTATGGAGTC | TGGTCGTGTG | AAGGATGTAA | GGCCTTTTTT | 360 |
| AAAAGAAGCA | TTCAAGGACA | TAATGATTAT | ATTTGTCCAG | СТАСАААТСА | GTGTACAATC | 420 |
| GATAAAAACC | GGCGCAAGAG | CTGCCAGGCC | TGCCGACTTC | GGAAGTGTTA | CGAAGTGGGA | 480 |
| ATGGTGAAGT | GTGGCTCCCG | GAGAGAGAGA | TGTGGGTACC | GCCTTGTGCG | GAGACAGAGA | 540 |
| AGTGCCGACG | AGCAGCTGCA | CTGTGCCGGC | AAGGCCAAGA | GAAGTGGCGG | CCACGCGCCC | 600 |
| CGAGTGCGGG | AGCTGCTGCT | GGACGCCCTG | AGCCCCGAGC | AGCTAGTGCT | CACCCTCCTG | 660 |
| GAGGCTGAGC | CGCCCCATGT | GCTGATCAGC | CGCCCCAGTG | CGCCCTTCAC | CGAGGCCTCC | 720 |
| ATGATGATGT | CCCTGACCAA | GTTGGCCGAC | AAGGAGTTGG | TACACATGAT | CAGCTGGGCC | 780 |
| AAGAAGATTC | CCGGCTTTGT | GGAGCTCAGC | CTGTTCGACC | AAGTGCGGCT | CTTGGAGAGC | 840 |
| TGTTGGATGG | AGGTGTTAAT | GATGGGGCTG | ATGTGGCGCT | CAATTGACCA | CCCCGGCAAG | 900 |
| CTCATCTTTG | CTCCAGATCT | TGTTCTGGAC | AGGGATGAGG | GGAAATGCGT | AGAAGGAATT | 960 |
| CTGGAAATCT | TTGACATGCT | CCTGGCAACT | ACTTCAAGGT | TTCGAGAGTT | AAAACTCCAA | 1020 |
| CACAAAGAAT | ATCTCTGTGT | CAAGGCCATG | ATCCTGCTCA | ATTCCAGTAT | GTACCCTCTG | 1080 |
| GTCACAGCGA | CCCAGGATGC | TGACAGCAGC | CGGAAGCTGG | CTCACTTGCT | GAACGCCGTG | 1140 |

ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG 1200

CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGTAA CAAGGGCATG 1260

GAACATCTGC TCAACATGAA GTGCAAAAAT GTGGTCCCAG TGTATGACCT GCTGCTGGAG 1320

ATGCTGAATG CCCACGTGCT TCGCGGGTGC AAGTCCTCCA TCACGGGGTC CGAGTGCAGC 1380

CCGGCAGAGG ACAGTAAAAG CAAAGAGGGC TCCCAGAACC CACAGTCTCA GTGA 1434

(2) INFORMATION FOR SEQ ID NO: 2:

5

10

25

30

35

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1251 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATGAATTACA GCATTCCCAG CAATGTCACT AACTTGGAAG GTGGGCCTGG TCGGCAGACC 60 ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGCACCTTT CTCCTTTAGT GGTCCATCGC 120 CAGTTATCAC ATCTGTATGC GGAACCTCAA AAGAGTCCCT GGTGTGAAGC AAGATCGCTA 180 GAACACCT TACCTGTAAA CAGAGAGACA CTGAAAAGGA AGGTTAGTGG GAACCGTTGC 240 GCCAGCCCTG TTACTGGTCC AGGTTCAAAG AGGGATGCTC ACTTCTGCGC TGTCTGCAGC 300 GATTACGCAT CGGGATATCA CTATGGAGTC TGGTCGTGTG AAGGATGTAA GGCCTTTTTT 360 AAAAGAAGCA TTCAAGGACA TAATGATTAT ATTTGTCCAG CTACAAATCA GTGTACAATC 420 GATAAAAACC GGCGCAAGAG CTGCCAGGCC TGCCGACTTC GGAAGTGTTA CGAAGTGGGA 480

ATGGTGAAGT GTGGCTCCCG GAGAGAGAGA TGTGGGTACC GCCTTGTGCG GAGACAGAGA 540 AGTGCCGACG AGCAGCTGCA CTGTGCCGGC AAGGCCAAGA GAAGTGGCGG CCACGCGCCC 600 CGAGTGCGGG AGCTGCTGCT GGACGCCCTG AGCCCCGAGC AGCTAGTGCT CACCCTCCTG 660 GAGGCTGAGC CGCCCCATGT GCTGATCAGC CGCCCCAGTG CGCCCTTCAC CGAGGCCTCC 720 ATGATGATGT CCCTGACCAA GTTGGCCGAC AAGGAGTTGG TACACATGAT CAGCTGGGCC 780 AAGAAGATTC CCGGCTTTGT GGAGCTCAGC CTGTTCGACC AAGTGCGGCT CTTGGAGAGC TGTTGGATGG AGGTGTTAAT GATGGGGCTG ATGTGGCGCT CAATTGACCA CCCCGGCAAG 900 CTCATCTTTG CTCCAGATCT TGTTCTGGAC AGGGATGAGG GGAAATGCGT AGAAGGAATT CTGGAAATCT TTGACATGCT CCTGGCAACT ACTTCAAGGT TTCGAGAGTT AAAACTCCAA 1020 CACAAAGAAT ATCTCTGTGT CAAGGCCATG ATCCTGCTCA ATTCCAGTAT GTACCCTCTG 1080 GTCACAGCGA CCCAGGATGC TGACAGCAGC CGGAAGCTGG CTCACTTGCT GAACGCCGTG 1140 ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG 1200 CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGGTG A 1251

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

35

25

30

5

| | (xi |) SEQI | SEQUENCE DESCRIPTION: SEQ ID NO: 3: | | | | | | | | | | | | | |
|----|-----------------|---------------|-------------------------------------|------------------------|-------------------------|---------------|----------------------|-----------|-----------|-----------|-----|-----------|-----------|-----------|-----------|-----|
| 5 | Су. 1 | s Ala | Val | Cys | Ser 5 | Asp | Tyr | Ala | Ser | Gly 10 | Tyr | His | Tyr | Gly | Val 15 | Trp |
| 3 | Se | r Cys | Glu | Gly 20 | Cys | Lys | Ala | Phe | Phe 25 | Lys | Arg | Ser | Ile | Gln 30 | Gly | His |
| 10 | Ası | n A sp | Tyr 35 | Ile | Cys | Pro | Ala | Thr 40 | Asn | Gln | Суѕ | Thr | Ile 45 | Asp | Lys | Asn |
| | Arc | g Arg 50 | Lys | Ser | Cys | Gln | Ala 55 | Cys | Arg | Leu | Arg | Lys 60 | Cys | Tyr | Glu | Val |
| 15 | G1 ₅ | y Met | | | | | | | | | | | | | | |
| | (2) INF | ORMAT: | ON I | FOR S | SEQ I | D N C | o: 4: | | | | | | | | | |
| 20 | (i) | (B) | LEI TYI | E CHANGTH: PE: 6 RANDI | : 233 amino EDNES | 3 ami caci | ino a id singl | cids | 5 | | | | | | | |
| 25 | (ii |) MOL | | | | | | | | | | | | | | |
| 30 | (xi |) SEQ | UENC | E DE: | SCRI | PTIO | N: Si | EQ II | D NO | : 4: | | | | | | |
| | Le 1 | u Val | Leu | Thr | Leu 5 | Leu | Glu | Ala | Glu | Pro 10 | Pro | His | Val | Leu | Ile 15 | Ser |
| 35 | Ar | g Pro | Ser | Ala 20 | Pro | Phe | Thr | Glu | Ala 25 | Ser | Met | Met | Met | Ser 30 | Leu | Thr |
| | Ly | s Leu | Ala 35 | Asp | Lys | Glu | Leu | Val 40 | His | Met | Ile | Ser | Trp | Ala | Lys | Lys |

| | Ile | Pro 50 | Gly | Phe | Val | Glu | Leu 55 | Ser | Leu | Phe | Asp | Gln 60 | Val | Arg | Leu | Leu |
|----|------------|------------|-------------------|------------|------------|------------|------------|------------|------------|------------|------------|--------------------|------------|--------------------|------------|------------|
| 5 | Glu 65 | Ser | Cys | Trp | Met | Glu 70 | Val | Leu | Met | Met | Gly 75 | Leu | Met | Trp | Arg | Ser 80 |
| 10 | Ile | Asp | His | Pro | Gly 85 | Lys | Leu | Ile | Phe | Ala 90 | Pro | Asp | Leu | Val | Leu 95 | Asp |
| | Arg | Asp | Glu | Gly 100 | Lys | Cys | Val | Glu | Gly 105 | Ile | Leu | Glu | Ile | Phe 110 | Asp | Met |
| 15 | Leu | Leu | Ala 115 | Thr | Thr | Ser | Arg | Phe 120 | Arg | Glu | Leu | Lys | Leu 125 | Gln | His | Lys |
| | Glu | Tyr 130 | Leu | Cys | Val | Lys | Ala 135 | Met | Ile | Leu | | A sn 140 | Ser | Ser | Met | Tyr |
| 20 | Pro 145 | Leu | Val | Thr | Ala | Thr 150 | Gln | Asp | Ala | Asp | Ser 155 | Ser | Arg | Lys | Leu | Ala 160 |
| 25 | His | Leu | Leu | Asn | Ala 165 | Val | Thr | Asp | Ala | Leu 170 | Val | Trp | Val | Ile | Ala 175 | Lys |
| | Ser | Gly | Ile | Ser 180 | Ser | Gln | Gln | Gln | Ser 185 | Met | Arg | Leu | Ala | As n 190 | Leu | Leu |
| 30 | Met | Leu | Leu 195 | Ser | His | Val | Arg | His 200 | Ala | Ser | Asn | Lys | Gly 205 | Met | Glu | His |
| | Leu | Leu 210 | Asn | Met | Lys | Cys | Lys 215 | Asn | Val | Val | | Va l 220 | Tyr | Asp | Leu | Leu |
| 35 | Leu 225 | Glu | Met | Leu | Asn | Ala | His | Val | Leu | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

| | | (A |) LE | NGTH | : 47 | 7 am | ino | acid | s | | | | | | | |
|----|----------------------|---|--|---|--|--|--|--|---|---|---|--|---|---|---|--|
| | | (B |) TY | PE: | amin | o ac | id | | | | | | | | | |
| | | (C |) ST | RAND | EDNE | ss: | sing | le | | | | | | | | |
| 5 | | (D |) ТО | POLO | GY: | unkn | own | | | | | | | | | |
| | (ii) | MOL | ECUL | Е ТҮ | PE: | prot | ein | | | | | | | | | |
| 10 | | | | | | | | | | | | | | | | |
| | (xi) | SEQ | UENC | E DE | SCRI | PTIO | N: S | EQ I | D NO | : 5: | | | | | | |
| | Met | Asn | Tyr | Ser | Ile | Pro | Ser | Asn | Val | Thr | Asn | Leu | Glu | Gly | Gly | Pro |
| | 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| 15 | | | | | | | | | | | | | | | | |
| | Gly | Arg | Gln | Thr 20 | Thr | Ser | Pro | Asn | Val 25 | Leu | Trp | Pro | Thr | Pro 30 | Gly | His |
| | Leu | Ser | Pro | Leu | Val | Val | His | Arg | Gln | Leu | Ser | His | Leu | Tvr | Ala | Glu |
| 20 | | | 35 | | | | | 40 | | | | | 45 | - 1 - | | 0.2.4 |
| | Pro | Gln 50 | Lys | Ser | Pro | Trp | Cys 55 | Glu | Ala | Arg | Ser | Leu 60 | Glu | His | Thr | Leu |
| 25 | Pro 65 | Val | Asn | Arg | Glu | Thr 70 | Leu | Lys | Arg | Lys | Val 75 | Ser | Gly | Asn | Arg | Cys 80 |
| | Ala | Ser | Pro | Val | Thr 85 | Gly | Pro | Gly | Ser | Lys 90 | Arg | Asp | Ala | His | Phe 95 | Cys |
| 30 | λla | Wa l | Cvc | 505 | 7 ~ ~ | Т | ת 1 ת | 505 | C1 | m | TT - | <i></i> | 61 | | _ | _ |
| | Ala | vai | Суѕ | 100 | Asp | TYL | ALA | ser | 105 | lyr | піз | Tyr | GTÀ | 110 | Trp | Ser |
| | Cys | Glu | Gly | Cys | Lys | Ala | Phe | Phe | Lys | Arg | Ser | Ile | Gln | Gly | His | Asn • |
| 35 | | | 115 | | | | | 120 | | | | | 125 | | | |
| | Asp | Tyr 130 | Ile | Суѕ | Pro | Ala | Thr 135 | Asn | Gln | Cys | Thr | Ile 140 | Asp | Lys | Asn | Arg |
| | 10 15 20 25 | (ii) (xi) Met 1 15 Gly Leu 20 Pro 25 Pro 65 Ala 30 Ala Cys | 10 (xi) SEQUENT SEQUEN | (B) TY (C) ST (D) TO | (B) TYPE: (C) STRAND (D) TOPOLO (ii) MOLECULE TY 10 (xi) SEQUENCE DE Met Asn Tyr Ser 1 15 Gly Arg Gln Thr 20 Leu Ser Pro Leu 35 Pro Gln Lys Ser 50 Pro Val Asn Arg 65 Ala Ser Pro Val 30 Ala Val Cys Ser 100 Cys Glu Gly Cys 115 Asp Tyr Ile Cys | (B) TYPE: amin (C) STRANDEDNE (D) TOPOLOGY: (ii) MOLECULE TYPE: 10 (xi) SEQUENCE DESCRI Met Asn Tyr Ser Ile 1 5 15 Gly Arg Gln Thr Thr 20 Leu Ser Pro Leu Val 35 Pro Gln Lys Ser Pro 50 25 Pro Val Asn Arg Glu 65 Ala Ser Pro Val Thr 85 30 Ala Val Cys Ser Asp 100 Cys Glu Gly Cys Lys 115 Asp Tyr Ile Cys Pro | (B) TYPE: amino ace (C) STRANDEDNESS: (D) TOPOLOGY: unkn (ii) MOLECULE TYPE: prot Met Asn Tyr Ser Ile Pro 1 5 Gly Arg Gln Thr Thr Ser 20 Leu Ser Pro Leu Val Val 35 Pro Gln Lys Ser Pro Trp 50 25 Pro Val Asn Arg Glu Thr 65 70 Ala Ser Pro Val Thr Gly 85 30 Ala Val Cys Ser Asp Tyr 100 Cys Glu Gly Cys Lys Ala 115 Asp Tyr Ile Cys Pro Ala | (B) TYPE: amino acid (C) STRANDEDNESS: sing (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: S Met Asn Tyr Ser Ile Pro Ser 1 5 Gly Arg Gln Thr Thr Ser Pro 20 Leu Ser Pro Leu Val Val His 35 Pro Gln Lys Ser Pro Trp Cys 50 Pro Val Asn Arg Glu Thr Leu 65 Pro Val Asn Arg Glu Thr Leu 65 Ala Ser Pro Val Thr Gly Pro 85 30 Ala Val Cys Ser Asp Tyr Ala 100 Cys Glu Gly Cys Lys Ala Phe 115 Asp Tyr Ile Cys Pro Ala Thr | (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ I. Met Asn Tyr Ser Ile Pro Ser Asn 1 5 Gly Arg Gln Thr Thr Ser Pro Asn 20 Leu Ser Pro Leu Val Val His Arg 35 40 Pro Gln Lys Ser Pro Trp Cys Glu 50 55 25 Pro Val Asn Arg Glu Thr Leu Lys 65 70 Ala Ser Pro Val Thr Gly Pro Gly 85 30 Ala Val Cys Ser Asp Tyr Ala Ser 100 Cys Glu Gly Cys Lys Ala Phe Phe 115 120 Asp Tyr Ile Cys Pro Ala Thr Asn | (C) STRANDEDNESS: single (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO Met Asn Tyr Ser Ile Pro Ser Asn Val 1 5 Gly Arg Gln Thr Thr Ser Pro Asn Val 20 25 Leu Ser Pro Leu Val Val His Arg Gln 35 70 Pro Gln Lys Ser Pro Trp Cys Glu Ala 50 55 Pro Val Asn Arg Glu Thr Leu Lys Arg 65 70 Ala Ser Pro Val Thr Gly Pro Gly Ser 85 Ala Val Cys Ser Asp Tyr Ala Ser Gly 100 105 Cys Glu Gly Cys Lys Ala Phe Phe Lys 115 120 Asp Tyr Ile Cys Pro Ala Thr Asn Gln | (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5: Met Asn Tyr Ser Ile Pro Ser Asn Val Thr 1 | (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5: Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn 1 5 10 (Si) Arg Gln Thr Thr Ser Pro Asn Val Leu Trp 20 25 Leu Ser Pro Leu Val Val His Arg Gln Leu Ser 35 Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser 50 55 Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val 65 70 75 Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg 85 90 Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His 100 105 Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser 115 120 Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr | (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5: Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu 1 5 10 (Sly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp Pro 20 25 Leu Ser Pro Leu Val Val His Arg Gln Leu Ser His 35 40 Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser Leu 50 55 60 25 Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val Ser 65 70 75 Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg Asp 85 90 30 Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr 100 105 Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile 35 115 120 Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile | (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5: Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu 1 | (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5: Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly 1 5 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5: Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly 1 5 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5: Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly 1 5 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5: Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly 10 20 25 30 Leu Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp Pro Thr Pro 20 25 40 45 Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser Leu Glu His 50 55 60 Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val Ser Gly Asn 65 70 75 Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg Asp Ala His 85 90 Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val 100 105 110 Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly 115 120 120 125 | (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5: Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly Gly Ileu Ser I |

| | Arg 145 | Lys | Ser | Cys | Gln | Ala 150 | Cys | Arg | Leu | Arg | Lys 155 | Cys | Tyr | Glu | Val | Gly 160 |
|--------|------------|------------|-------------------|------------|------------|------------|------------|------------|------------|----------------|------------|------------|------------|-------------------|------------|------------|
| | | | | | | | | | | | | | | | | |
| 5 | Met | Val | Lys | Cys | Gly 165 | Ser | Arg | Arg | Glu | Arg 170 | Cys | Gly | Tyr | Arg | Leu 175 | Val |
| | Arg | Arg | Gln | Arg 180 | Ser | Ala | Asp | Glu | Gln 185 | Leu | His | Cys | Ala | Gly 190 | Lys | Ala |
| 10 | Lys | Arg | Ser 195 | Gly | Gly | His | Ala | Pro 200 | Arg | Val | Arg | Glu | Leu 205 | Leu | Leu | Asp |
| | Ala | Leu 210 | Ser | Pro | Glu | Gln | Leu 215 | Val | Leu | Thr | Leu | Leu 220 | Glu | Ala | Glu | Pro |
| 15 | Pro 225 | His | Val | Leu | Ile | Ser 230 | Arg | Pro | Ser | Ala | Pro 235 | Phe | Thr | Glu | Ala | Ser 240 |
| 20 | Met | Met | Met | Ser | Leu 245 | Thr | Lys | Leu | Ala | Asp 250 | Lys | Glu | Leu | Val | His 255 | Met |
| | Ile | Ser | Trp | Ala 260 | Lys | Lys | Ile | Pro | Gly 265 | Phe | Val | Glu | Leu | Ser 270 | Leu | Phe |
| 25 | Asp | Gln | Val 275 | Arg | Leu | Leu | Glu | Ser 280 | Cys | Trp | Met | Glu | Val 285 | Leu | Met | Met |
| 20 | Gly | Leu 290 | Met | Trp | Arg | Ser | Ile 295 | Asp | His | Pro | Gly | Lys 300 | Leu | Ile | Phe | Ala |
| 30 | Pro 305 | Asp | Leu | Val | Leu | Asp 310 | Arg | Asp | Glu | Gly | Lys 315 | Cys | Val | Glu | Gly | Ile 320 |
| 35 | Leu | Glu | Ile | Phe | Asp 325 | Met | Leu | Leu | Ala | Thr 330 | Thr | Ser | Arg | Phe | Arg 335 | Glu |
| | Leu | Lys | Leu | Gln 340 | | Lys | Glu | Tyr | Leu 345 | Cys | Val | Lys | Ala | Met 350 | Ile | Leu |

Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser Gln Gln Ser Met Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Ser Asn Lys Gly Met Glu His Leu Leu Asn Met Lys Cys Lys Asn Val Val Pro Val Tyr Asp Leu Leu Glu Met Leu Asn Ala His Val Leu Arg Gly Cys Lys Ser Ser Ile Thr Gly Ser Glu Cys Ser Pro Ala Glu Asp . 20 Ser Lys Ser Lys Glu Gly Ser Gln Asn Pro Gln Ser Gln (2) INFORMATION FOR SEQ ID NO: 6: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 416 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn Leu Glu Gly Gly Pro

Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp Pro Thr Pro Gly His Leu Ser Pro Leu Val Val His Arg Gln Leu Ser His Leu Tyr Ala Glu Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser Leu Glu His Thr Leu Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val Ser Gly Asn Arg Cys Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg Asp Ala His Phe Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly Met Val Lys Cys Gly Ser Arg Arg Glu Arg Cys Gly Tyr Arg Leu Val Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His Cys Ala Gly Lys Ala Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg Glu Leu Leu Asp Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro

| | | | 210 | | | | | 215 | | | | | 220 | | | | |
|---|-----|------------|------------|------------|------------|---------------------|--------------------|------------|------------|-------------|----------------|------------|------------|------------|-------------------|------------|------------|
| | | Pro 225 | His | Val | Leu | Ile | Ser 230 | Arg | Pro | Ser | Ala | Pro 235 | Phe | Thr | Glu | Ala | Ser 240 |
| | 5 | Met | Met | Met | Ser | Leu 2 4 5 | Thr | Lys | Leu | Ala | Asp 250 | Lys | Glu | Leu | Val | His 255 | Met |
| | 10 | Ile | Ser | Trp | Ala 260 | Lys | Lys | Ile | Pro | Gly 265 | Phe | Val | Glu | Leu | Ser 270 | Leu | Phe |
| . A. e.g. | | Asp | Gln | Val 275 | Arg | Leu | Leu | Glu | Ser 280 | Cys | Trp | Met | Glu | Val 285 | Leu | Met | Met |
| 1 97 2 30 3 40 3 40 3 40 3 40 3 40 3 40 3 40 3 | 15 | Gly | Leu 290 | Met | Trp | Arg | Ser | Ile 295 | Asp | His | Pro | Gly | Lys 300 | Leu | Ile | Phe | Ala |
| | 0.0 | Pro 305 | Asp | Leu | Val | Leu | As p 310 | Arg | Asp | Glu | Gly | Lys 315 | Cys | Val | Glu | Gly | Ile 320 |
| 2 (27) | 20 | Leu | Glu | Ile | Phe | Asp 325 | Met | Leu | Leu | Ala | Thr 330 | Thr | Ser | Arg | Phe | Arg 335 | Glu |
| 1.5. 1. | 25 | Leu | Lys | Leu | Gln 340 | His | Lys | Glu | Tyr | Leu 345 | Суѕ | Val | Lys | Ala | Met 350 | Ile | Leu |
| | | Leu | Asn | Ser 355 | Ser | Met | Tyr | Pro | Leu 360 | Va l | Thr | Ala | Thr | Gln 365 | Asp | Ala | Asp |
| | 30 | Ser | Ser 370 | Arg | Lys | Leu | Ala | His 375 | Leu | Leu | Asn | Ala | Val 380 | Thr | Asp | Ala | Leu |
| | 2.5 | Val 385 | Trp | Val | Ile | Ala | Lys 390 | Ser | Gly | Ile | Ser | Ser 395 | Gln | Gln | Gln | Ser | Met 400 |
| | 35 | Arg | Leu | Ala | Asn | Leu 405 | Leu | Met | Leu | Leu | Ser 410 | His | Val | Arg | His | Ala 415 | Arg |

,

| | (2) INFORMATION FOR SEQ ID NO: 7: | |
|-----|--|----|
| | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 29 base pairs | |
| 5 | (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: both | |
| | (D) TOPOLOGY: unknown | |
| 10 | (ii) MOLECULE TYPE: cDNA | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7: | |
| 15 | GGIGAYGARG CWTCIGGITG YCAYTAYGG | 29 |
| | (2) INFORMATION FOR SEQ ID NO: 8: | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| 20 | (A) LENGTH: 29 base pairs | |
| | (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: single | |
| | (D) TOPOLOGY: linear | |
| 25 | (ii) MOLECULE TYPE: cDNA | |
| 2.0 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8: | |
| 30 | AAGCCTGGSA YICKYTTIGC CCAIYTIAT | 29 |
| | (2) INFORMATION FOR SEQ ID NO: 9: | ٠, |
| 35 | (i) SEQUENCE CHARACTERISTICS: | • |
| | (A) LENGTH: 22 base pairs | |
| | (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: single | |

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

| 5 | | |
|----|---|----|
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9: | |
| | TGTTACGAAG TGGGAATGGT GA | 22 |
| 10 | (2) INFORMATION FOR SEQ ID NO: 10: | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 24 base pairs | |
| | (B) TYPE: nucleic acid | |
| 15 | (C) STRANDEDNESS: single | |
| | (D) TOPOLOGY: linear | |
| | (ii) MOLECULE TYPE: cDNA | |
| 20 | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10: | |
| | | |
| | TTGACACCAG ACCAACTGGT AATG | 24 |
| 25 | (2) INFORMATION FOR SEQ ID NO: 11: | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 24 base pairs | |
| 30 | (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: single | |
| | (D) TOPOLOGY: linear | |
| | (ii) MOLECULE TYPE: cDNA | • |
| 35 | | |
| | | |

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

| | GGTGGCGACG ACTCCTGGAG CCCG | 24 |
|-----|---|----|
| | (2) INFORMATION FOR SEQ ID NO: 12: | |
| 5 | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 22 base pairs | |
| | (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: single | |
| | (D) TOPOLOGY: linear | |
| 10 | (ii) MOLECULE TYPE: cDNA | |
| | (, | |
| | | |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12: | |
| | GTACACTGAT TTGTAGCTGG AC | 22 |
| | (2) INFORMATION FOR SEQ ID NO: 13: | |
| 20 | | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 20 base pairs | |
| | (B) TYPE: nucleic acid | |
| 2.5 | (C) STRANDEDNESS: single | |
| 25 | (D) TOPOLOGY: linear | |
| | (ii) MOLECULE TYPE: cDNA | |
| | | |
| 30 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13: | |
| | CCATGATGAT GTCCCTGACC | 20 |
| | | * |
| 35 | (2) INFORMATION FOR SEQ ID NO: 14: | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 20 base pairs | |
| | (B) TYPE: nucleic acid | |

(C) STRANDEDNESS: single

| | (D) TOPOLOGY: linear | |
|----|--|----|
| 5 | (ii) MOLECULE TYPE: cDNA | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14: | |
| 10 | TCGCATGCCT GACGTGGGAC | 20 |
| | (2) INFORMATION FOR SEQ ID NO: 15: | |
| 15 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 24 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear | |
| 20 | (ii) MOLECULE TYPE: cDNA | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15: | |
| 25 | GGCSTCCAGC ATCTCCAGSA RCAG | 24 |
| | (2) INFORMATION FOR SEQ ID NO: 16: | |
| 30 | (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 20 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear | - |
| 35 | (iii) MOLEGULE TYPE PNA | |
| | (ii) MOLECULE TYPE: cDNA | |

١.

| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16: | |
|----|---|---|
| | GGAAGCTGGC TCACTTGCTG | 20 |
| 5 | (2) INFORMATION FOR SEQ ID NO: 17: | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | | |
| | | |
| 10 | | |
| | (D) TOPOLOGY: linear | |
| | (ii) MOLECULE TYPE: cDNA | |
| 15 | | |
| 10 | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17: | |
| | | |
| | TCTTGTTCTG GACAGGGATG | 20 |
| 20 | (0) 7) 70 70 70 70 70 70 70 70 70 70 70 70 70 | |
| | (2) INFORMATION FOR SEQ ID NO: 18: | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | (A) LENGTH: 20 base pairs | |
| 25 | (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: single | |
| | (D) TOPOLOGY: linear | |
| | (ii) MOLECULE TYPE: cDNA | |
| 30 | | |
| | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18: | |
| | | ٠, |
| 35 | GCATGGAACA TCTGCTCAAC | 20 |
| | (2) INFORMATION FOR SEQ ID NO: 19: | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| | 10 15 20 25 | GGAAGCTGGC TCACTTGCTG 5 (2) INFORMATION FOR SEQ ID NO: 17: (1) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: cDNA 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17: TCTTGTTCTG GACAGGGATG 20 (2) INFORMATION FOR SEQ ID NO: 18: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: cDNA 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18: 35 GCATGGAACA TCTGCTCAAC (2) INFORMATION FOR SEQ ID NO: 19: |

(A) LENGTH: 21 base pairs

| | (B) TYPE: nucleic acid | |
|----|--|------|
| | (C) STRANDEDNESS: single | |
| | (D) TOPOLOGY: linear | |
| 5 | | |
| | (ii) MOLECULE TYPE: cDNA | |
| | | |
| | | |
| | | |
| 10 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19: | |
| | | |
| | AGCAAGTTCA GCCTGTTAAG T | 21 |
| | | |
| | (2) INFORMATION FOR SEQ ID NO: 20: | |
| 15 | | |
| | (i) SEQUENCE CHARACTERISTICS: | |
| 1 | (A) LENGTH: 1257 base pairs | |
| | (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: double | |
| 20 | (D) TOPOLOGY: linear | |
| | (iii) NOTEGULE WARE ARE | |
| | (ii) MOLECULE TYPE: cDNA | |
| | | |
| 25 | | |
| 20 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20: | |
| | , and a second of the second o | |
| | ATGAATTACA GCATTCCCAG CAATGTCACT AACTTGGAAG GTGGGCCTGG TCGGCAGACC | 60 |
| | | |
| 30 | ACAAGCCCAA ATGTGTTGTG GCCAACACCT GGGCACCTTT CTCCTTTAGT GGTCCATCGC | 120 |
| | | |
| | CAGTTATCAC ATCTGTATGC GGAACCTCAA AAGAGTCCCT GGTGTGAAGC AAGATCGCTA | 180 |
| | | |
| | GAACACACCT TACCTGTAAA CAGAGAGACA CTGAAAAGGA AGGTTAGTGG GAACCGTTGC | ,240 |
| 35 | | |
| | GCCAGCCCTG TTACTGGTCC AGGTTCAAAG AGGGATGCTC ACTTCTGCGC TGTCTGCAGC | 300 |
| | | |
| | GATTACGCAT CGGGATATCA CTATGGAGTC TGGTCGTGTG AAGGATGTAA GGCCTTTTTT | 360 |
| | | |

AAAAGAAGCA TTCAAGGACA TAATGATTAT ATTTGTCCAG CTACAAATCA GTGTACAATC 420 GATAAAAACC GGCGCAAGAG CTGCCAGGCC TGCCGACTTC GGAAGTGTTA CGAAGTGGGA 480 ATGGTGAAGT GTGGCTCCCG GAGAGAGAGA TGTGGGTACC GCCTTGTGCG GAGACAGAGA 540 AGTGCCGACG AGCAGCTGCA CTGTGCCGGC AAGGCCAAGA GAAGTGGCGG CCACGCGCCC 600 CGAGTGCGGG AGCTGCTGCT GGACGCCCTG AGCCCCGAGC AGCTAGTGCT CACCCTCTG 660 GAGGCTGAGC CGCCCCATGT GCTGATCAGC CGCCCCAGTG CGCCCTTCAC CGAGGCCTCC 720 ATGATGATGT CCCTGACCAA GTTGGCCGAC AAGGAGTTGG TACACATGAT CAGCTGGGCC 780 AAGAAGATTC CCGGCTTTGT GGAGCTCAGC CTGTTCGACC AAGTGCGGCT CTTGGAGAGC 840 TGTTGGATGG AGGTGTTAAT GATGGGGCTG ATGTGGCGCT CAATTGACCA CCCCGGCAAG 900 CTCATCTTTG CTCCAGATCT TGTTCTGGAC AGGGATGAGG GGAAATGCGT AGAAGGAATT 960 CTGGAAATCT TTGACATGCT CCTGGCAACT ACTTCAAGGT TTCGAGAGTT AAAACTCCAA 1020 CACAAAGAAT ATCTCTGTGT CAAGGCCATG ATCCTGCTCA ATTCCAGTAT GTACCCTCTG 1080 GTCACAGCGA CCCAGGATGC TGACAGCAGC CGGAAGCTGG CTCACTTGCT GAACGCCGTG 1140 ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG 1200 CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGGTC TGCCTGA 1257

(2) INFORMATION FOR SEQ ID NO: 21:

5

10

15

20

25

30

35

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 418 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

 $\sqrt{}$

| | (xi) | SEQ | UENC | E DE | SCRI | PTIO | N: S | EQ I | D NO | : 21 | : | | | | | |
|------------------------|----------|-------|------|------------|-----------|----------|------|------|------------|--------------|--------------|-----|-----|------------|------------|------|
| 5 | • | | _ | | | _ | | | _ | | | | | | | |
| | Met 1 | : Asn | Tyr | Ser | 11e 5 | Pro | Ser | Asn | Val | Thr 10 | Asn | Leu | Glu | Gly | Gly 15 | Pro |
| | C). | | G1 - | m | m) | G | | | **. 1 | - | _ | | m1 | • | a 1 | |
| 10 | GI | / Arg | GIN | 20 | Thr | ser | Pro | Asn | 25 | Leu | Trp | Pro | Thr | 30 | GIĄ | His |
| | Lei | ı Ser | Pro | Leu | Val | Val | His | Arg | Gln | Leu | Ser | His | Leu | Tvr | Ala | Glu |
| | | | 35 | | | | | 40 | | | | | 45 | 4 | | |
| 15 | Pro | Gln | Lys | Ser | Pro | Trp | Cys | Glu | Ala | Arg | Ser | Leu | Glu | His | Thr | Leu |
| 2 | | 50 | | | | | 55 | | | | | 60 | | | | |
| 1 mm | | Val | Asn | Arg | Glu | | Leu | Lys | Arg | Lys | Val | Ser | Gly | Asn | Arg | |
| 20 | 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| 27 2 ⁷ 2 | Ala | Ser | Pro | Val | Thr 85 | Gly | Pro | Gly | Ser | Lys 90 | Arg | Asp | Ala | His | Phe 95 | Cys |
| - 27 19 10 10 | | | | | 03 | | | | | 30 | | | | | 93 | |
| 25 | Ala | Val | Суѕ | Ser 100 | Asp | Tyr | Ala | Ser | Gly 105 | Tyr | His | Tyr | Gly | Val 110 | Trp | Ser |
| | Cyc | Glu | Cly | Cuc | Tuc | λla | Dho | Dh o | T | N m m | 505 | Tla | Cl= | C1 | Uia | 2 |
| | C y S | Giu | 115 | СуЗ | БУЗ | Ala | rne | 120 | гуэ | ALG | 261 | 116 | 125 | GIY | птэ | ASII |
| 30 | Asp | Tyr | Ile | Cys | Pro | Ala | Thr | Asn | Gln | C y s | Thr | Ile | Asp | Lys | Asn | Arg |
| | | 130 | | | | | 135 | | | | | 140 | | | | |
| | | Lys | Ser | Cys | Gln | | Cys | Arg | Leu | Arg | | Cys | Tyr | Glu | Val | Gly |
| 35 | 145 | • | | | | 150 | | | | | 1 5 5 | | | | | 160 |
| | Met | : Val | Lys | Cys | _ | Ser | Arg | Arg | Glu | _ | Cys | Gly | Tyr | Arg | | Val |
| | | | | | 165 | | | | | 170 | | | | | 175 | |
| | Arc | Ara | Gln | Ara | Ser | Ala | Asp | Glu | Gln | Leu | His | Cvs | Ala | Glv | Lvs | Ala |

Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg Glu Leu Leu Leu Asp Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val Glu Leu Ser Leu Phe Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met Glu Val Leu Met Met Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly Lys Leu Ile Phe Ala Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys Cys Val Glu Gly Ile Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val Lys Ala Met Ile Leu Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu

Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser Gln Gln Gln Ser Met

- 49 -

ŗ

385 390 395 400 Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Arg 405 410 415 5 Ser Ala (2) INFORMATION FOR SEQ ID NO: 22: 10 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 34 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 15 (ii) MOLECULE TYPE: cDNA (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22: CTTGGATCCA TAGCCCTGCT GTGATGAATT ACAG 34 (2) INFORMATION FOR SEQ ID NO: 23: 25 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 33 base pairs (B) TYPE: nucleic acid 30 (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: cDNA 35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GATGGATCCT CACCTCAGGG CCAGGCGTCA CTG

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1898 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

5

10

15

20

25

30

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CACGAATCTT TGAGAACATT ATAATGACCT TTGTGCCTCT TCTTGCAAGG TGTTTTCTCA 60 GCTGTTATCT CAAGACATGG ATATAAAAAA CTCACCATCT AGCCTTAATT CTCCTTCCTC 120 CTACAACTGC AGTCAATCCA TCTTACCCCT GGAGCACGGC TCCATATACA TACCTTCCTC 180 CTATGTAGAC AGCCACCATG AATATCCAGC CATGACATTC TATAGCCCTG CTGTGATGAA 240 TTACAGCATT CCCAGCAATG TCACTAACTT GGAAGGTGGG CCTGGTCGGC AGACCACAAG 300 CCCAAATGTG TTGTGGCCAA CACCTGGGCA CCTTTCTCCT TTAGTGGTCC ATCGCCAGTT 360 ATCACATCTG TATGCGGAAC CTCAAAAGAG TCCCTGGTGT GAAGCAAGAT CGCTAGAACA 420 CACCTTACCT GTAAACAGAG AGACACTGAA AAGGAAGGTT AGTGGGAACC GTTGCGCCAG 480 CCCTGTTACT GGTCCAGGTT CAAAGAGGGA TGCTCACTTC TGCGCTGTCT GCAGCGATTA 540 CGCATCGGGA TATCACTATG GAGTCTGGTC GTGTGAAGGA TGTAAGGCCT TTTTTAAAAG 600 AAGCATTCAA GGACATAATG ATTATATTTG TCCAGCTACA AATCAGTGTA CAATCGATAA 660 AAACCGGCGC AAGAGCTGCC AGGCCTGCCG ACTTCGGAAG TGTTACGAAG TGGGAATGGT 720

GAAGTGTGGC TCCCGGAGAG AGAGATGTGG GTACCGCCTT GTGCGGAGAC AGAGAAGTGC 780 CGACGAGCAG CTGCACTGTG CCGGCAAGGC CAAGAGAAGT GGCGGCCACG CGCCCGAGT 840 GCGGGAGCTG CTGCTGGACG CCCTGAGCCC CGAGCAGCTA GTGCTCACCC TCCTGGAGGC 900 TGAGCCGCCC CATGTGCTGA TCAGCCGCCC CAGTGCGCCC TTCACCGAGG CCTCCATGAT 960 GATGTCCCTG ACCAAGTTGG CCGACAAGGA GTTGGTACAC ATGATCAGCT GGGCCAAGAA 1020 GATTCCCGGC TTTGTGGAGC TCAGCCTGTT CGACCAAGTG CGGCTCTTGG AGAGCTGTTG 1080 GATGGAGGTG TTAATGATGG GGCTGATGTG GCGCTCAATT GACCACCCCG GCAAGCTCAT CTTTGCTCCA GATCTTGTTC TGGACAGGGA TGAGGGGAAA TGCGTAGAAG GAATTCTGGA 1200 AATCTTTGAC ATGCTCCTGG CAACTACTTC AAGGTTTCGA GAGTTAAAAC TCCAACACAA 1260 AGAATATCTC TGTGTCAAGG CCATGATCCT GCTCAATTCC AGTATGTACC CTCTGGTCAC 1320 AGCGACCCAG GATGCTGACA GCAGCCGGAA GCTGGCTCAC TTGCTGAACG CCGTGACCGA 1380 TGCTTTGGTT TGGGTGATTG CCAAGAGCGG CATCTCCTCC CAGCAGCAAT CCATGCGCCT 1440 GGCTAACCTC CTGATGCTCC TGTCCCACGT CAGGCATGCG AGTAACAAGG GCATGGAACA 1500 TCTGCTCAAC ATGAAGTGCA AAAATGTGGT CCCAGTGTAT GACCTGCTGC TGGAGATGCT 1560 GAATGCCCAC GTGCTTCGCG GGTGCAAGTC CTCCATCACG GGGTCCGAGT GCAGCCCGGC 1620 AGAGGACAGT AAAAGCAAAG AGGGCTCCCA GAACCCACAG TCTCAGTGAC GCCTGGCCCT 1680 GAGGTGAACT GGCCCACAGA GGTCACAAGC TGAAGCGTGA ACTCCAGTGT GTCAGGAGCC 1740 TGGGCTTCAT CTTTCTGCTG TGTGGTCCCT CATTTGGTGA TGGCAGGCTT GGTCATGTAC 1800 CATCCTTCCC TCCACCTTCC CAACTCTCAG GAGTCGGTGT GAGGAAGCCA TAGTTTCCCT 1860

10

1.5

20

25

30

TGTTAGCAGA GGGACATTTG AATCGAGCGT TTCCACAC

1898

| (2) | INFORMATION | FOR | SEO | ΙD | NO: | 25: |
|-----|-------------|-----|-----|----|-----|-----|
| | | | | | | |

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 530 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Met Asp Ile Lys Asn Ser Pro Ser Ser Leu Asn Ser Pro Ser Ser Tyr

1 5 10 15

Asn Cys Ser Gln Ser Ile Leu Pro Leu Glu His Gly Ser Ile Tyr Ile
20 25 30

Pro Ser Ser Tyr Val Asp Ser His His Glu Tyr Pro Ala Met Thr Phe
35 40 45

Tyr Ser Pro Ala Val Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn 50 55 60

Leu Glu Gly Gly Pro Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp 65 70 75 80

Pro Thr Pro Gly His Leu Ser Pro Leu Val Val His Arg Gln Leu Ser 85 90 95

His Leu Tyr Ala Glu Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser
100 105 110

Leu Glu His Thr Leu Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val

15

10

5

20

25

30

Ser Gly Asn Arg Cys Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg Asp Ala His Phe Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His Asn Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val Gly Met Val Lys Cys Gly Ser Arg Arg Glu Arg Cys Gly Tyr Arg Leu Val Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His Cys Ala Gly Lys Ala Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg Glu Leu Leu Asp Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val

Glu Leu Ser Leu Phe Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met

325 330 335

| Glu | Val | Leu | Met 340 | Met | Gly | Leu | Met | Trp 345 | Arg | Ser | Ile | Asp | His 350 | Pro | Gly |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--------------------|----------------|------------|------------|------------|
| Lys | Leu | Ile 355 | Phe | Ala | Pro | Asp | Leu 360 | Val | Leu | Asp | Arg | Asp 365 | Glu | Gly | Lys |
| Cys | Val 370 | Glu | Gly | Ile | Leu | Glu 375 | Ile | Phe | Asp | Met | Leu 380 | Leu | Ala | Thr | Thr |
| Ser 385 | Arg | Phe | Arg | Glu | Leu 390 | Lys | Leu | Gln | His | Lys 395 | Glu | Tyr | Leu | Cys | Val 400 |
| Lys | Ala | Met | Ile | Leu 405 | Leu | Asn | Ser | Ser | Met 410 | Tyr | Pro | Leu | Val | Thr 415 | Ala |
| Thr | Gln | Asp | Ala 420 | Asp | Ser | Ser | Arg | Lys 425 | Leu | Ala | His | Leu | Leu 430 | Asn | Ala |
| Val | Thr | Asp 435 | Ala | Leu | Val | Тгр | Val 440 | Ile | Ala | Lys | Ser | Gly 445 | Ile | Ser | Ser |
| Gln | Gln 450 | Gln | Ser | Met | Arg | Leu 455 | Ala | Asn | Leu | Leu | M et 460 | Leu | Leu | Ser | His |
| Val 465 | Arg | His | Ala | Ser | Asn 470 | Lys | Gly | Met | Glu | His 475 | Leu | Leu | Asn | Met | Lys 480 |
| Cys | Lys | Asn | Val | Val 485 | Pro | Val | Tyr | Asp | Leu 490 | Leu | Leu | Glu | Met | Leu 495 | Asn |
| Ala | His | Val | Leu 500 | Arg | Gly | Суѕ | Lys | Ser 505 | Ser | Ile | Thr | Gly | Ser 510 | Glu | Cys |

Ser Pro Ala Glu Asp Ser Lys Ser Lys Glu Gly Ser Gln Asn Pro Gln

Ser Gln

530

| | | (2) INFORMATION FOR SEQ ID NO: 26: | |
|-------|----|--|----|
| | 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 30 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear | |
| | 10 | (ii) MOLECULE TYPE: other nucleic acid | |
| | 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26: | |
| | | GTGCGGATCC TCTCAAGACA TGGATATAAA | 30 |
| | 20 | (2) INFORMATION FOR SEQ ID NO: 27: | |
| Tong. | 25 | (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 25 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear | |
| | | (ii) MOLECULE TYPE: other nucleic acid | |
| | 30 | | |
| | | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27: | |
| | 35 | AGTAACAGGG CTGGCGCAAC GGTTC | 25 |
| | | (2) INFORMATION FOR SEQ ID NO: 28: | |

(i) SEQUENCE CHARACTERISTICS:

- 56 -

(A) LENGTH: 22 base pairs(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

Ĺ

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

ACTGGCGATG GACCACTAAA GG

22

15

10